Docket No.: 3516P007
Express Mail No.: EL651820845US

UNITED STATES PATENT APPLICATION

FOR

METHOD AND SYSTEM TO DETERMINE THE GEOGRAPHIC LOCATION OF A NETWORK USER

Inventor:

Thomas H. Hampton

Prepared By:

BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN LLP
12400 Wilshire Boulevard, 7th Floor
Los Angeles, California 90025-1026
(310) 207-3800

METHOD AND SYSTEM TO DETERMINE THE GEOGRAPHIC LOCATION OF A NETWORK USER

BACKGROUND

Field of the Invention

[001] This invention relates to the field of networked computers. More specifically, it relates to a system and method to determine the geographic location of a network user based on that user's current network (*e.g.*, Internet Protocol (IP)) address and previously stored mapping request information obtained from related network addresses.

Background

[002] The Internet and personal computers have become ubiquitous in modern society. People regularly access the Internet via personal computers and other devices for any number of purposes. Traditionally, most home and small business computer users have accessed the Internet via a modem plugged in to a telephone line that connects an Internet user to an Internet service provider (ISP). It is through the ISP that the Internet user receives an IP address and accesses the Internet.

[003] At any given moment, each Internet host is uniquely identified by an Internet Protocol (IP) address and a domain name. The format of an IP address is a 32 bit numeric address which is expressed as four integer numbers, each separated by a dot, commonly referred to as dotted-decimal notation. Each number must be in the range from zero or one to 255; the first byte cannot be zero. For example, 209.1.25.249 is a valid IP address. It is the IP address that is used by underlying network software to identify a personal computer and other network hosts that communicate over the Internet.

[004] Though the Internet Protocol has absolute requirements regarding the nature of the dotted decimal notation and how specific addresses become associated with specific pieces of network equipment, as a matter of practice, the organizations responsible for their allocation do so in substantial blocks of addresses. For example, when an ISP requests additional address space, the ISP normally receives a contiguous range of monotonically ascending addresses, rather than a totally random assortment of unrelated addresses. This administrative mechanism leads to the following result:

addresses that differ only in the least significant, forth byte of the 32 bit address are almost always on the same local area network, and consequently, are physically located within a relatively short distance, such as a mile or so, from one another.

[005] Prior art techniques exist that rely on indirect means to essentially guess the physical location of network devices based upon their IP address. For example, a utility called "whois" allows an Internet user to find out information about domain names and IP addresses including to whom the domain name and IP address is registered. Another utility called "tracert" allows an Internet user to trace the route of a data packet from the user to a remote host specified by the user. Yet another utility called "nslookup" allows an Internet user to look up the IP address of a domain name. In addition, there is a tool that provides the latitude and longitude of the geographic location of an Internet host. This tool, called "ip2ll," converts an IP address to latitude and longitude based on "whois" database information. More specifically, the city of the administrative contact for the IP address is obtained from a "whois" database and then converted into latitude and longitude. The address of the administrative contact is the mailing address of a party responsible for registration of the IP address and any associated domain name. This address may be geographically distant from the equipment over which the party has control, since one person can efficiently serve as the contact for nationally or globally distributed network equipment. Information regarding "ip2ll" is available from the Digital Computer Laboratory, Department of Computer Science, University of Illinois at Urbana-Champaign, 1304 West Springfield Avenue, Urbana, Illinois. In contrast, network users are almost always physically located near network equipment accessed by local telephone calls to an ISP, as long distance telephone calls are expensive. No tools are currently available that allow an Internet user to be associated with a defined geographical area.

BRIEF SUMMARY OF THE INVENTION

[006] A system and method to determine the geographic location of a network user based on a plurality of network addresses and a plurality of mapping request information. In one embodiment, the method involves receiving a plurality of mapping requests from a plurality of network users

identified by a plurality of network addresses. A geographic location of a network user associated with a network address is then derived based on the plurality of network addresses and the plurality of mapping requests. In one embodiment, the network address is an Internet Protocol (IP) address. The method may be implemented, in one embodiment, as part of a system that executes as software on a computer.

BRIEF DESCRIPTION OF THE DRAWINGS

[007] **Figure 1** depicts a network environment in which one embodiment of the method and system for determining the geographic location of a network user operates.

[008] **Figure 2** depicts a flow diagram of one embodiment of a process for determining the geographic location of a network user.

DETAILED DESCRIPTION

[009] A method and system for determining the geographic location of a network user is described. In one embodiment, the method involves making determinations of the geographic location of network users based on previously stored mapping request information from the same network address or a similar network address via a simple predictive heuristic. A confidence factor may be assigned to the result. The method also involves assembling a large number of passively collected observations about the affinity of network addresses for various locations. A newly or previously encountered network address is matched by the heuristic with previously observed behavior to arrive at the probable physical location of a network user identified by a network address.

[0010] **Figure 1** depicts a network environment in which one embodiment of the method and system for determining the geographic location of a network user operates. In one embodiment, the process for determining the geographic location of a network user may be implemented in software that is stored on and runs in a mapping server computer 110. In other embodiments the process may be implemented in hardware or a combination of hardware and software. Mapping server computer 110 may be any server computer. In one embodiment, server computer 110 includes processor 154, memory 156, and communications interface 158 each of which is coupled to bus 152. Storage device 118 is also coupled to bus 152. Storage device 118 may be any readable and writeable non-volatile machine readable medium including, for example, magnetic media such as a hard disk drive, electronic media such a flash memory, optical memory such as a compact disk memory device, or any device that includes a machine readable medium.

[0011] In one embodiment, mapping server computer 110 may include mapping software 112 that processes mapping requests, analyzer software 114 that identifies geographic locations of network users, and database software 116 that may be used to store network user network address data and corresponding mapping requests issued by the network user. The database software may be any database, including for example, databases that provide support for the structured query language (SQL). In one embodiment,

mapping software 112, analyzer software 114, and database software 116 are stored on a hard disk drive version of storage device 118. In one embodiment, the process for determining the geographic location of a network user is achieved by executing processor 154 using storage device 118 and memory 156 to execute mapping software 112, analyzer software 114, and database software 116. Although only single instances of the components of mapping server 110 are depicted, multiple processors, memory devices, storage devices, and other components, not shown, may be included in mapping server 110.

[0012] Mapping server computer 110 may be coupled to a network such as the Internet 120 or other global wide area network via communications interface 158. Communications interface 158 may be any device or adapter that allows to communication over a network such as the Internet, including, modems, network cards, and the like. Multiple personal computing devices 130 are coupled to the Internet and submit mapping requests which may be serviced by the mapping software on the mapping server 110. Personal computing devices 130 may be personal computers, personal digital assistants (PDAs), cellular telephones, computer tablets, or other computing devices capable of communicating over the Internet. As with mapping server 110, the computing devices each include a processor, memory, bus, communications interface, storage device, and other similar components.

[0013] Although mapping server computer 110 is depicted as a single computer, in another embodiment, multiple servers in a cluster, local area network (LAN), subgroup, etc. may be used. In these embodiments, the multiple servers may include dedicated graphics servers, database servers, user interface servers, etc. and may be separated from the Internet or other network by a firewall, gateway, etc.

[0014] In one embodiment, the mapping software on the mapping server receives and satisfies numerous mapping requests received from personal computing devices and other network hosts, including other servers, not shown. In addition to responding to the mapping requests, the mapping software stores the mapping request data and corresponding network addresses via database software. This may be achieved according to well-known database methods via database software located on the mapping server

or a dedicated database server included in a multiple server LAN, subgroup, etc.

[0015] The analyzer software determines the geographic location of network users based on the mapping request data and corresponding network addresses stored in the database. In one embodiment, the analyzer software delimits the entire world, or a portion thereof, into defined geographic regions and populates each of the regions with network addresses such as Internet Protocol (IP) addresses based on mapping requests issued from the IP address. That is, for each mapping request falling in a defined geographic region, the network address of the network user issuing the mapping request is associated with the particular geographic region in which the mapping request falls. Because of the way IP addresses have been allocated historically, IP addresses tend to cluster within a particular geographical region. The analyzer software may then determine which clusters or groupings of network addresses may be associated with particular geographic regions based on those geographic regions having the majority of occurrences of the particular network address.

[0016] In one embodiment, a common portion of an IP address is used in the analysis. In this embodiment, the first three numbers or bytes comprising 32 bit IP addresses may be used as the common portion of the IP address that is used in this analysis rather than the entire network address. In this embodiment, the correspondence between network addresses and mapping requests is analyzed based only on the first three bytes of the network address rather than the entire network address. In this way, groups or subgroups of addresses having a common portion may be found to predominate in certain geographic regions. That is, the techniques described herein exploits the fact that IP address that differ only in the forth byte of the 32 bit address are almost always on the same local area network. Consequently, addresses that differ only in the forth byte of the 32 bit address are generally physically located within a mile or so of one another. As discussed above, this is a result of the fact that IP addresses have been allocated in blocks of sequentially grouped addresses.

[0017] In one embodiment, the defined geographic region may be a 50 mile square area. However, various sizes of areas and geometric shapes may be used to define the geographic regions in various embodiments. In one

embodiment, the geographic regions may be uniform in size. In another embodiment, the geographic regions may vary in size and/or shape based on the proximity to major cities, whether a region is densely populated, whether a region is sparsely populated, etc. In one embodiment, regions may be defined by population rather than area such that each region may be comprised of the same defined population such as 100,000 persons, 500,000 persons, 1,000,000 persons, etc.

[0018] A third party computer/server 140 which is also coupled to the Internet may request that the mapping server provide a network address to geographic location translation. The mapping server 110 may then respond with the requested geographic location pursuant to the techniques described herein. The geographic location may be specified as a latitude and longitude pair that may be the center of the particular geographic region, may be a geocode, or may be any other geographic or regional designation. The third party may then use the location information for any desired purpose, such as to display geographically relevant information to the Internet user including advertisements, a weather forecast pertinent to the user, local events of interest to the user, etc. In one embodiment, the mapping server may also include a content server component or other server component such that in response to receiving a request for a network address to geographic translation from a third party, a geographically relevant advertisement, weather report, events listing, or other information is also provided to the third party requestor, or may be provided directly to the user pursuant to instructions from the third party requestor.

[0019] **Figure 2** depicts a flow diagram of one embodiment of a process for determining the geographic location of a network user. The process is performed by processing logic that my comprise hardware, software, or a combination of both. A mapping server receives a mapping request from an Internet user that implicitly includes the IP address of the Internet user, as shown in block 210. The mapping server then stores mapping request data and the corresponding IP address of the Internet user, as shown in block 220. The mapping server services the mapping request, as shown in block 230. The details of servicing the mapping request are not described in detail herein so as not to obscure the present invention. A cycle of receiving, storing and

servicing mapping requests as shown in blocks 210, 220 and 230 may proceed for many iterations. A check is then made to determine whether enough data has been stored, as shown in block 240. Whether there is enough stored data is based on a determination of whether a statistically sufficient sample has been stored. This determination may be made according to any well-known methods. Until there has been enough data stored so as to be able to statistically analyze multiple mapping requests and corresponding IP addresses, the mapping server continues to receive, store and service mapping requests, as shown in blocks 210, 220 and 230.

[0020] When enough data has been stored, the mapping server analyzes the correspondence between multiple IP addresses and multiple mapping requests to determine the geographic locations of the multiple IP addresses, as shown in block 250. This analysis and determination may occur regularly after a certain minimum amount of data sufficient for analysis is received, and may be constantly refined based on each mapping request received. In this way, the analysis constitutes an ongoing learning process. In one embodiment, only the most recent mapping requests and IP addresses are used. This helps to accommodate for temporary IP address assignment. In one embodiment, the most recent mapping requests and corresponding IP addresses may include all information received over a period of, for example, four days, one week, one month, three months, etc. Enough data is maintained so as to be able to provide an accurate result while also accounting for stale or no longer valid data resulting from the re-use of temporary IP addresses.

[0021] In one embodiment, database software may be used to maintain and analyze the pertinent mapping request data and corresponding IP addresses. In one embodiment, analyzer software may access the database software to obtain information needed to analyze or otherwise extrapolate or evaluate the correspondence between multiple IP addresses and multiple mapping requests. The analyzer software determines which geographic regions are associated with particular IP addresses or ranges of IP addresses in a simple, straight-forward manner. The analyzer software assigns IP addresses to those geographic regions having the majority of mapping requests that originate from Internet users having the same or similar IP addresses. In one embodiment, the IP addresses are considered to be the same or similar when

the IP addresses share the first three bytes of the dotted decimal version of the IP addresses. This may be referred to as the common portion of the IP addresses.

[0022] The mapping server may then receive a request to provide the geographic location of an Internet user identified by an IP address, as shown in block 260. The mapping server then provides the geographic location of the specified Internet user identified by an IP address based on the analyzed correspondence between the multiple stored IP addresses and corresponding mapping requests, as shown in block 270. In one embodiment, the geographic location provided by the mapping server may also be accompanied by an accuracy rating or confidence factor so that the recipient of the geographic location of the specified Internet user may better use the information provided. For example, the accuracy rating/confidence factor may be a percentage, a number from 1 to 10, a number from 1 to 5, a group of letters of the alphabet, or any other sliding scale that may be used to express an ordered relationship. The rating may reflect whether a strong or weak number of IP addresses make up a majority of the IP addresses associated with a geographic region. For example, if 10,000 mapping requests have been stored all of which have the same or similar first three bytes of IP address from originating users, and 5,000 of the mapping requests are for a first region, 3,000 of the mapping requests are for a second region, 1,000 requests are for a third region and 500 mapping requests are each for fourth and fifth regions, then a response to an IP address to location translation for an IP address having the same or similar IP address may include the first region and may include a high rating. The high rating may be appropriate as 50% of the mapping requests from the particular family of IP addresses fell into the particular region. The high rating may be understood contextually if the typical distribution of mapping requests for IP addresses results in a broad distribution among more than 10 regions such that no region typically has more than one-third of the mapping requests for a particular IP address range. This is only an example, and the accuracy and tuning of the system will vary over time, among geographic regions, and according to other factors.

[0023] In the foregoing specification, the invention has been described with reference to specific embodiments thereof. It will, however, be evident

that various modifications and changes can be made thereto without departing from the broader spirit and scope of the invention as set forth in the appended claims. The specification and drawings are, accordingly, to be regarded in an illustrative rather than a restrictive sense.